CSE574 Introduction to Machine Learning

Support Vector Machine

Jue Guo

University at Buffalo

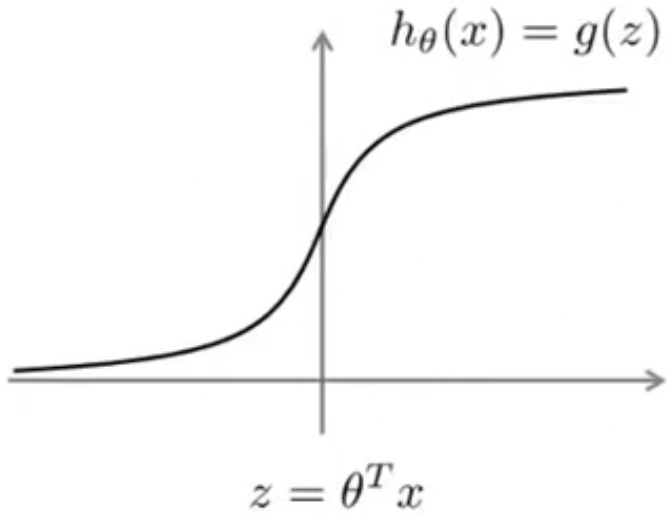
January 28, 2024

Outline

Alternative View of Logistic Regression

A quick review: $h_\theta(x) = \frac{1}{1+e^{-\theta^T x}}$

- if $y = 1$, we want $h_\theta(x) \approx 1$, $\theta^T x \gg 0$
- if $y = 0$, we want $h_\theta(x) \approx 0$, $\theta^T x \ll 0$



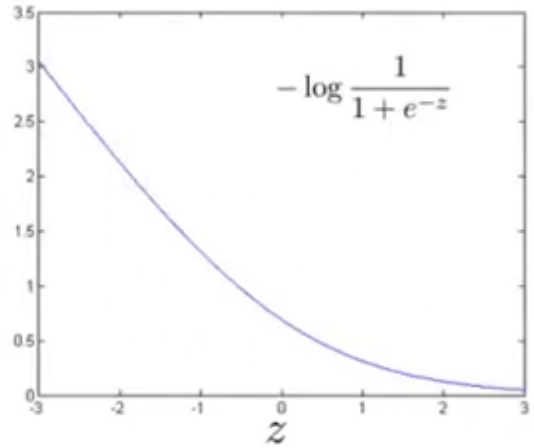
The cost of a single example:

$$-\left(y \log h_\theta(x) + (1-y) \log\left(1 - h_\theta(x)\right)\right)$$

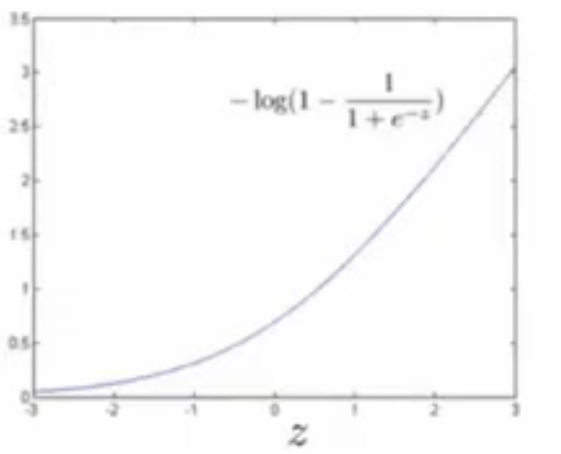
$$= -\, y \log \frac{1}{1+e^{-\theta^T x}} - (1-y) \log\left(1 - \frac{1}{1+e^{-\theta^T x}}\right)$$

$$-y \log \frac{1}{1+e^{-\theta^T x}} - (1-y) \log \left(1 - \frac{1}{1+e^{-\theta^T x}}\right)$$

if $y = 1$ (want $\theta^T x \gg 0$)



if $y = 0$ (want $\theta^T x \ll 0$)



Cost Function of Logistic Regression

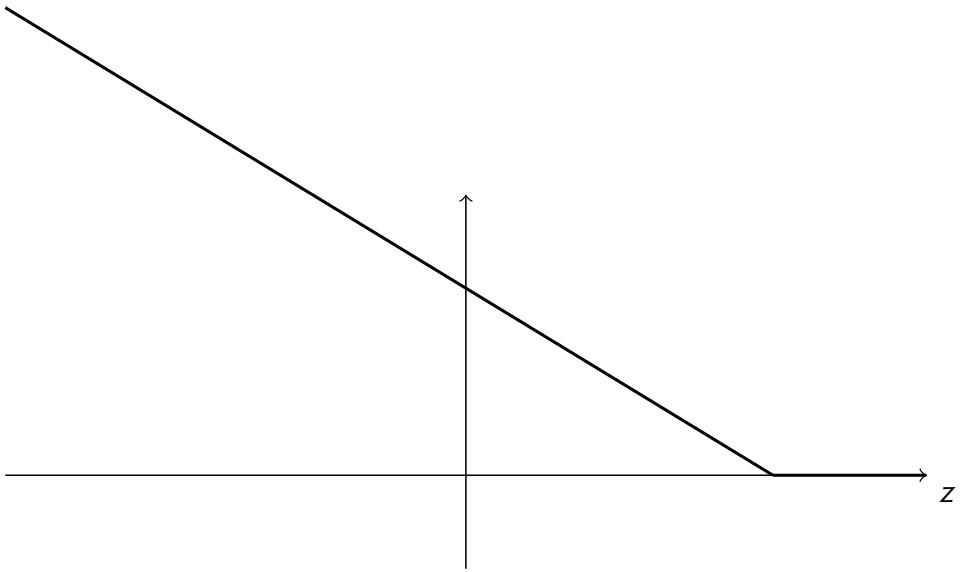
$$\min_{\theta} \frac{1}{m} \left[\sum_{i=1}^{m} y^{(i)} \left(-\log h_\theta \left(x^{(i)} \right) \right) + \left(1 - y^{(i)} \right) \left(-\log \left(1 - h_\theta \left(x^{(i)} \right) \right) \right) \right] + \frac{\lambda}{2m} \sum_{j=1}^{n} \theta_j^2$$

Cost Function of Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} \text{cost}_1 \left(\theta^T x^{(i)} \right) + \left(1 - y^{(i)} \right) \text{cost}_0 \left(\theta^T x^{(i)} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$

Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} \text{cost}_1 \left(\theta^T x^{(i)} \right) + \left(1 - y^{(i)} \right) \text{cost}_0 \left(\theta^T x^{(i)} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_j^2$$



Questions?